Open Science Grid Consortium and other Joint Grid Projects

Ruth Pordes

Drivers & Scope

Drivers: Effective and efficient support for Fermilab stakeholder computing needs.

- Significant use of resources remote from Fermilab part of Run II roadmap and clearly part of (US)-CMS computing models.
- Sharing of resources and policies to increase overall effectiveness of ensemble of computing.
- Facilitates inclusive and coherent approach to enabling use of opportunistically available resources at Universities and other Labs.

Scope: Leadership contributions to a National Common Grid Infrastructure which is Open.

- Aligned and synergistic with the US CMS data grid.
- Presenting resources to this common grid infrastructure
- Provide assured use of resources agreed upon through agreements (MOUs). For CMS that is at minimum within US CMS & CMS; for Fermilab this is with the Laboratory experiment and user base.
- Support opportunistic use of resources by organizations outside of US CMS and Fermilab.
- Contribute to the common infrastructure services, within the scope of the existing missions of US CMS and Fermilab CD, by leveraging developments and deployments with OSG mission.
  - Joint development projects;
  - Common storage and compute facilities,
  - Intra-Lab grid: FermiGrid,
  - US CMS projects (US CMS funding 3 CMS-OSG positions in CD.)
  - DOE-SciDAC funding of SRM & PPDG

Services

Experiment (US CMS, Run II, SDSS, Minos, etc)
simulation, production and analyses running on OSG.

Data management - ie. SamGrid, CMS data management.

Workflow management ie. MOP, Jim, MCRunjob, D0Runjob..

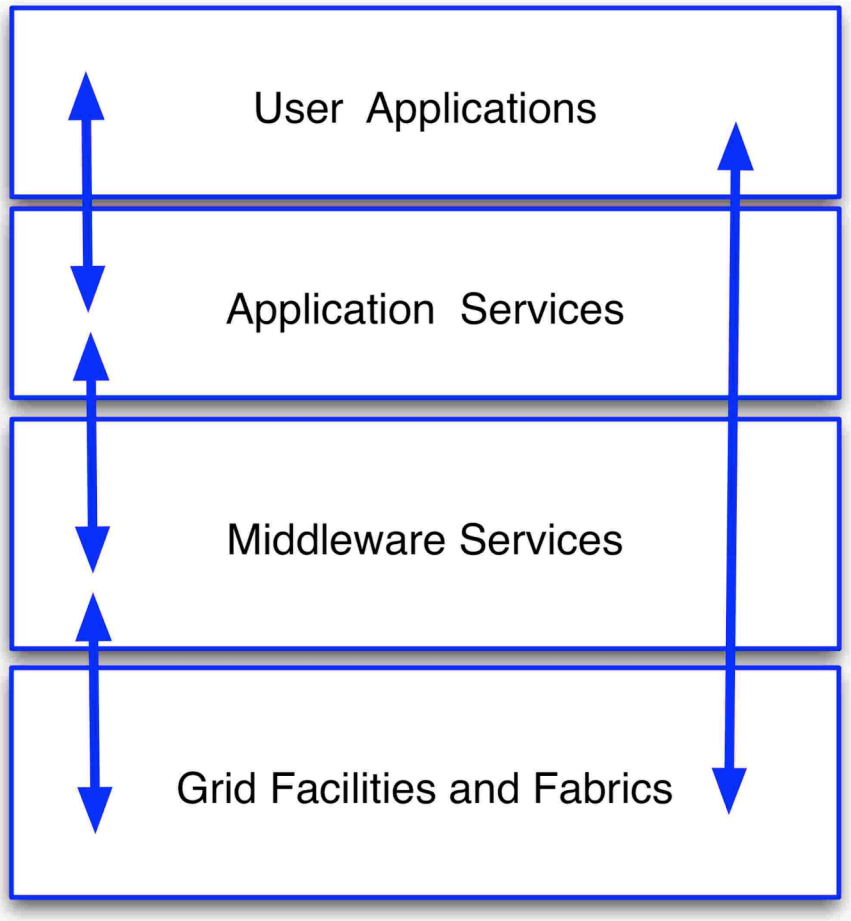
Experiment monitoring & accounting ie. Jim-MIS, Monalisa, +VO-accounting project,

Information, user & role management- ie. VOX, VO-Privilege, +extensions +information services.

Storage systems - Upper Storage; Lower Storage

Batch systems & compute farms - +sandboxing, +diagnostic tools,

Managed networks - Lambdastation, WAN researcg

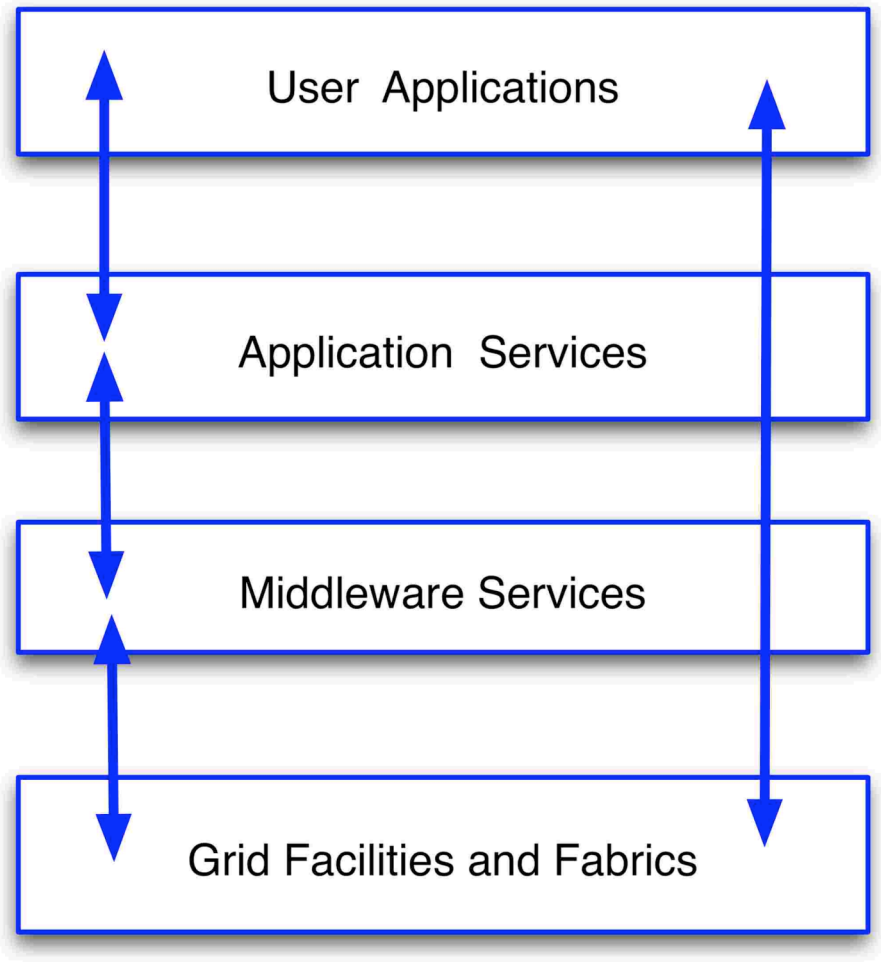


Interfaces

- Service Indexes - UDDI and/or Web service Discovery
- Execution Environment - ?
- Job Descriptions and History - ?
- Data Placement and Tracking- ?

- Job Scheduling - Class-Ads/Condor-G
- Data Movement - GridFTP
- Storage Management - SRM
- Job Execution - Gram

- Site Policies and Permissions - Authz callout;
- Monitoring, Information, Accounting - ?



End to End Systems

- Open Science Grid Release 0.2 running applications with mulit-user, multi-role and multi-organization assured and opportunistic use of compute and storage services.
  - Including problem detection and diagnostics; Heartbeats etc.
- Grid Operations
  - Support processes across multiple “helpdesks” and support organizations. Including Fermilab Customer Support, US CMS, VDT etc.
  - Demonstrated Security Incident Handling and Response.
- Interoperability across multiple implementations of
  - Compute Services: done in Grid3
  - Storage Services: SRM/dCache; LBNL DRM. Tests coming from PPDG common project; Upper Storage etc.
  - Grids: FermiGrid, LCG, TeraGrid, Glow.
- Define and meet performance and scaling metrics.
  - n terabytes transferred daily, for many days, across CMS Tier-1/Tier-2 sites
  - Nx1000 job floods sustained over days.

OSG Contributions

- Grid Services developed at Fermilab and in US CMS are architected and designed to help define and demonstrate the common interfaces and with attention to the roadmap. Integrated into development projects - many of which are presented today:
  e.g. Upper Storage: Storage Management (SRM), Authorization, Monitoring, Accounting interfaces to storage services.
- FermiGrid plan includes interfacing to and support use from OSG.
- Fermilab CD and PPDG contributing "staff office" resources:
  - PPDG Technical Coordinator, Dane Skow, & PPDG Common project specifically contributing deliverables to OSG. PPDG Executive team sponsoring Blueprint, Application meeting.
  - OSG web pages and administrative infrstructure.
- Fermilab (CD and US CMS) stepping up to lead and coordinate:
  - TG-storage - Rob; TG-Monitoring - Conrad; TG-SupportCenters - Adam; TG-Policy - Markus. TG-Security - Dane; TG-Networks(?) - Don
- US CMS Tier-2 proposals include commitments to and leverage of OSG roadmap and interfacing.

Milestones

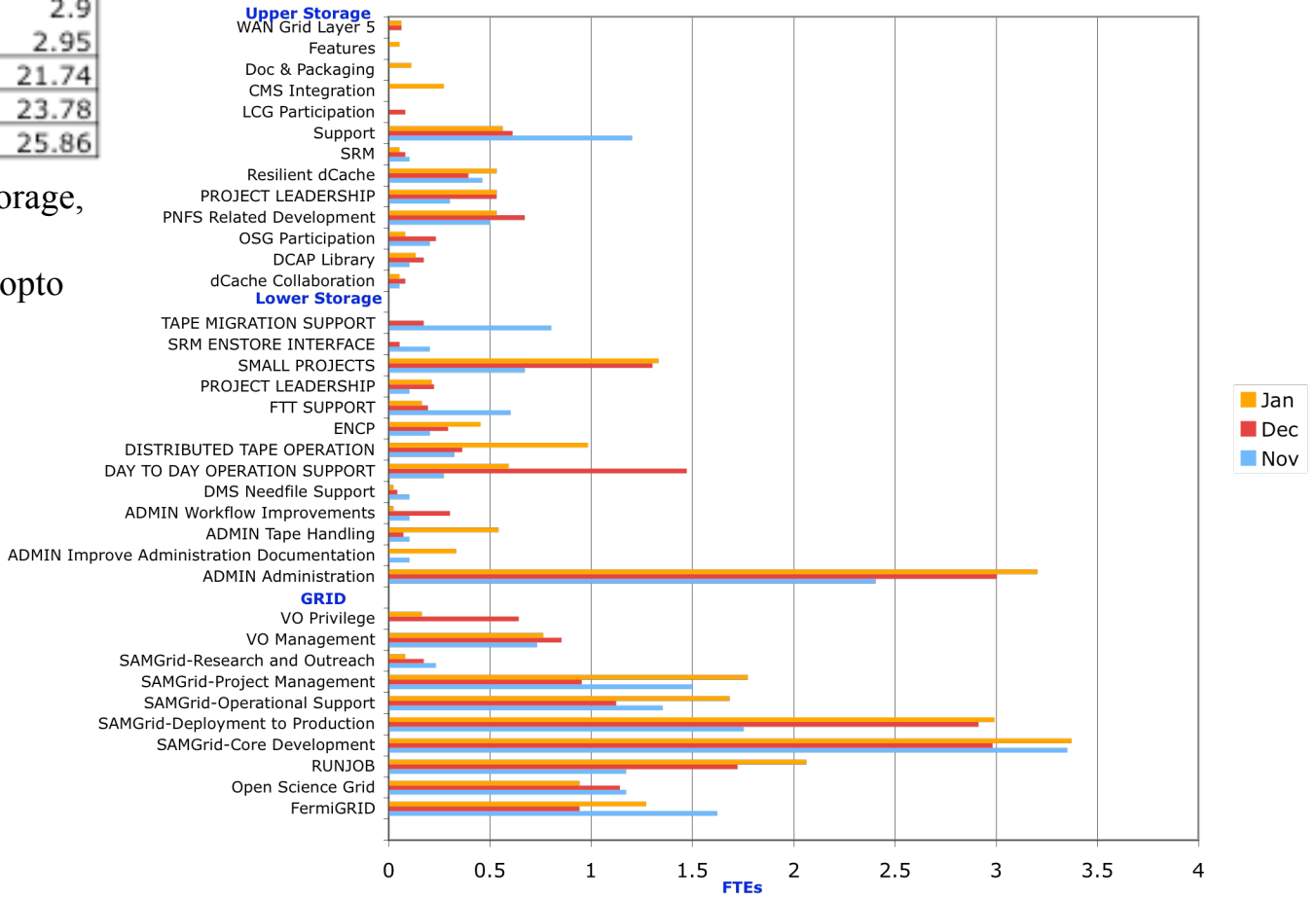
- OSG Deployment, release 0.2 of the core infrastructure: Spring 2005
  - Integration fest at UofC 15-17th February
- OSG operations of persistent infrastructure following this.
- D0 and CDF reprocessing & increasingly remote analyses.
- OSG Release 0.3 in Fall 2005 with new milestones every 6 months with added capabilities, scale and performance.
- For US CMS grid services and infrastructure (OSG?) must deliver to experiment's needs:
  - Tier-1/Tier-2 distributed analysis v0 in 2005.
  - Support for DC06 data challenge - end of 2005
  - Support for start of data taking - end of 2006.

Effort Profile

Category	Data	Total
GRID	Sum of Nov	12.87
	Sum of Dec	13.42
	Sum of Jan	15.08
Lower Storage	Sum of Nov	5.96
	Sum of Dec	7.46
	Sum of Jan	7.83
Upper Storage	Sum of Nov	2.91
	Sum of Dec	2.9
	Sum of Jan	2.95
Total Sum of Nov		21.74
Total Sum of Dec		23.78
Total Sum of Jan		25.86

Does not include US CMS:~1 FTE Storage, ~1 FTE VO services.

Nov before redistribution of vacation/opto



Needs and Risks

- Several areas really need attention & there is insufficient effort available:
  - Network and storage management, integration and access
  - Accounting and Auditing
  - Integration, Validation and sustaining of the infrastructure
  - Ongoing Interoperability between LCG and OSG
  - Information Services and Monitoring
  - Execution environment, sandboxing and diagnostics (troubleshooting)
- Can we sustain the commitment of the stakeholders and contributors for the several years to make a common vision viable.
  - Ian Foster in his article for the German escience magazine "Open Science Grid is a grass-roots effort"(?)
- Insufficiently mature infrastructure or effort to support Run II production needs.
- Schedule needs of US CMS (and US ATLAS) may put too much pressure on delivering parochial solutions.
- We allow too much extraneous noise in the system and don't get the necessary intellectual focus.
- Current hires are cancelled and common scope must be curtailed.